Int. J. Dev. Biol. 53: 371-381 (2009)
doi: 10.1387/ijdb.082664za

THE INTERNATIONAL JOURNAL OF DEVELOPMENTAL BIOLOGY
www.intjdevbiol.com

Evolution and pleiotropy of TRITHORAX function in *Arabidopsis*

ZOYA AVRAMOVA*

School of Biological Sciences, University of Nebraska, Lincoln, NE, USA

ABSTRACT The *SET* domain-containing genes of the *TRITHORAX* family encode epigenetic factors that maintain the expression of targeted genes. *Trithorax* homologs have been found in both animals and plants. Since these are thought to have evolved multicellularity independently, common mechanisms of epigenetic regulation must be evolutionarily ancient and derived from a common ancestor. In addition, each lineage has evolved unique mechanisms to expand the original repertoire of epigenetic functions. Phylogenetic analysis of SET domain proteins has outlined some intriguing evolutionary trends. In plants, epigenetic gene silencing mechanisms have been aggressively pursued. In contrast, studies of epigenetic mechanisms maintaining active gene expression have been scarce. The goal of this review is to draw attention to this gap. Trithorax function in plants are analyzed here in an evolutionary context tracing phylogenetic relationships between the histone methyltransferase activities in unicellular and multicellular domains of life. The involvement of two members of the *Arabidopsis* Trithorax family, *ARABIDOPSIS HOMOLOG of TRITHORAX1 (ATX1),* and *ARABIDOPSIS HOMOLOG of TRITHORAX2 (ATX2),* in developmental and adaptation processes of the plant is overviewed.

KEY WORDS: *epigenetics, trithorax, Arabidopsis, evolution, histone modifications*

Introduction

The commonly found definition of epigenetics is that of a *"study of heritable changes in genome function that occur without a change in DNA sequence"* (Bird, 2007 and ref. therein). Epigenetic mechanisms regulate a broad spectrum of processes including development, differentiation, embryonic stem cells maintenance, senescence, disease and cancer (rev. in Kiefer, 2007; Kouzarides, 2007 and ref. therein). Unable to fully silence expressed genes or to activate completely silent genes, epigenetic regulators maintain established states. Their ability to propagate information of active/repressed gene states from mother-to daughter cells has defined them as bearers of the "cell memory" (Pirrotta, 1998). However, growing evidence that neuronal gene-expression states are also regulated by epigenetic mechanisms, despite evidence that neuronal cells do not divide, has opened space for a broader unifying definition that keeps "the sense of prevailing usage but avoids constraints imposed by stringently required heritability" (Bird, 2007). Epigenetic events might reflect "the structural adaptation of chromosomal regions so as to register, signal, or perpetuate altered activity states" (Bird, 2007).

Known epigenetic regulators operate within a system composed of at least three different molecular mechanisms: DNA methylation, chromatin modifications and RNA-based mechanisms, representing the "three pillars of epigenetics" (Grant-Downton and Dickinson, 2005; 2006). Epigenetic (non-Mendelian) events have been recognized in plants as well: variegated gene expression in *Oenothera blandina* after X-ray chromosomal disruptions and translocations (Catcheside, 1938, 1949) is similar to PEV in *Drosophila*; paramutation (Brink, 1950; Chandler and Stam, 2004), somaclonal variation (Kaeppler *et al*., 2000), nucle-

Abbreviations used in this paper: AG, agamous; AP, apetala; ATX, Arabidopsis homolog of trithorax; CLF, curly leaf; DAST, domain associated with SET in Trithorax (a substitute name for FYRN-FYRC); E(z), enhansor of zeste; FLC, flowering locus C; FYRC, phenylalanine-tyrosine rich at the C-terminus; FYRN, phenylalanine-tyrosine rich at the N-terminus; JA, jasmonic acid; PcG, polycomb group; PRC, polycomb repressive complex; SA, salicylic acid; SET, Su(var)3-9, E(z), trithorax; Su(z)12, suppressor of zeste 12; SWN, swing; TrxG, trithorax group; VRN, vernalyzation independent.

***Address correspondence to:** Zoya Avramova. SBS, Beadle Center, UNL, Lincon, NE 68588, USA. Fax: +1-402-472-8722. e-mail: zavramova2@unlnotes.unl.edu

Dedication: This paper is dedicated to the memory of Roumen Tsanev, a versatile scholar, an inspiring teacher and a pioneer researcher of chromatin structure and epigenetics. (For interview with Tsanev, see Nonchev and Tsaneva, 2009, doi: 10.1387/ijdb.082679sn).

Published online: 24 April 2009.

ISSN: Online 1696-3547, Print 0214-6282

olar dominance (Pikaard and Pontes, 2007, and ref.therein), and transgene silencing (Napoli *et al*., 1990) are epigenetic phenomena driven by molecular mechanisms similar to those operating in animals (for recent reviews see Grant-Downton and Dickinson, 2005; 2006; Pikaard and Pontes, 2007).

The Polycomb group (PcG) and Trithorax group (TrxG) complexes have been among the best studied paradigms of epigenetic mechanisms. Expression states of homeotic genes (active or silent) are established early in development but are maintained and faithfully propagated throughout cellular divisions by the counteracting activities of PcG/TrxG complexes (Grimaud *et al*., 2006, and ref. therein). In contrast to animals, organ development in plants is not restricted to the embryonic stage: the lateral organs (leaves), the reproductive organs (flowers), and the seeds originate from the same undifferentiated meristem active throughout the life cycle. Because differentiation and organogenesis are not fixed in embryogenesis, it was not evident that PcG/TrxG functions would participate in plant developmental processes. However, the discovery that genes encoding PcG/TrxG homologs play roles in development and survival strategies of *Arabidopsis* changed dramatically this view. In plants, as in animals, development of a wrong organ at a wrong place (homeosis) is a consequence of a mutation of a homeotic gene. Unlike the animal counterparts, however, the plant homeotic genes are not clustered and belong to the MADS-box family of transcription factors.

Two major classes of PcG repressor complexes, PRC2 and PRC1, mediate formation of transcription-resistant chromatin structure at the animal *Hox* genes (Grimaud *et al*., 2006, and ref. therein). Plant cells are totipotent and, accordingly, plants have developed epigenetic mechanisms that are related, although not identical, with those used by animals or yeasts (Avramova, 2002; Loidl, 2004). For example, PRC2 complexes of both animal and plant origin are conserved structurally and functionally, in terms of histone methyltransferase activity (Table 1; Chanvivattana *et al*., 2004; Schubert *et al*., 2005); however, plant PRC1 homologs have not been identified. Plant epigenetic silencing mechanisms have been extensively studied (rev. in Pien and Grossniklaus, 2007; Henderson and Jacobsen, 2007; Ng *et al*., 2007; Reyes, 2006; Schubert *et al*., 2005). In contrast, studies of the plant TRX homologs have been lagging behind. Given the tightly balanced PcG/TrxG interaction for the control of homeotic gene, it is logical to expect that counteracting H3K27 and H3K4-modifying activities would be regulating plant genes as well. Genome-wide analysis has reveled that the H3K27me3 labels are distributed at single-gene regions (Turck *et al*., 2007; Zhang *et al*., 2007a); whole-genome H3K4me3 distribution has not been reported. Unless more attention is shifted towards elucidating the role of the Trithorax component of the PcG/TrxG mechanism, our understanding of epigenetic processes in plants might remain severely unbalanced.

Evolution of *SET* domain genes and histone methylation marks across unicellular and multicellular kingdoms

The highly conserved (~150 amino acids) SET peptide is found in a number of proteins belonging to both repressive [Su(var)3-9, E(z)] and activating [Trithorax and Ash1] chromatin complexes (Stassen *et al*., 1995). The genes encoding SET-domain proteins are ancient, existing in the Bacterial Domain of life, but have proliferated and evolved novel functions linked with the appearance of eukaryotes (Alvarez-Venegas *et al*., 2007b). Because SET-domain proteins can modify chromatin by methylating specific lysines on the histone tails (Rea *et al*., 2000), it is not surprising that *SET*-domain genes were found in all eukaryotes, from the simple unicellular to the multicellular animals and plants. Whole-genome comparative analyses have revealed a positive correlation between genome evolution, overall gene content, organismal complexity, and functional requirements linked with the appearance of multicellularity (Hedges *et al*., 2004). It was suggested that core biological functions, common for both unicellular and multicellular organisms, would be carried out by a comparable number of orthologous proteins, while specialized processes unique to multicellulars would use novel proteins (Aravind and Subramanian, 1999).

SET domain proteins have intrinsic preference for specific histone lysine-residues; a methylation sign at a particular lysine may have enormous consequences for the transcriptional competence of pertinent genes (Kouzarides, 2007). Intuitively, it is expected that the numbers, types, and biochemical specificity of SET domain proteins, as well as the patterns and complexity of the epigenetic marks 'written' by them, would reflect the occurrence of novel functions. Indeed, involvement of *SET*-domain genes in multicellular functions like proliferation, ontogenesis, adhesion-mediated silencing, and disease (Sparmann and van Lohuzien, 2006) correlates with increased numbers and *SET*-domain family types in the genomes of animals and plants (Alvarez-Venegas and Avramova, 2002; Veerappan *et al.,* 2008).

Comparisons across the unicellular and multicellular domains of life, however, are complicated by lack of knowledge of the ancestral *SET* domain genes at the branching point. The monophyletic fungal group (*Ascomycetes*) offers a simplified model to trace the evolution of *SET*-domain genes in an evolutionarily well-defined group containing unicellular and multicellular members. It includes the unicellular yeasts (hemiascomycetes, *Saccharomycetes,* and archiascomycetes, *S. pombe*), as well as

TABLE 1

THE PCG/TRXG COMPLEXES IN THE THREE KINGDOMS OF LIFE

Complexes		Components		Function
PcG				
	animals	**plants**	**yeast**	
	Pc			Establishes and maintains
PRC1	Psc	-	-	repressive states*
	Ph			
	Ring			
	E (z)	CLF/SWN/MEA	-	
	Su(z)12	EMF/VRN2/FIS2	-	
PRC2				Establishes H3K27 me3*
	Esc	FIE	-	
	P55	MSI1	MSI1	
TrxG				
	Core complex	**(not isolated)**	***Compass***	
	MLL 1-4	ATX1-5	SET1	
	WDR5	AT3G49660	CPS30	Establishes H3K4 me3*
	RbBP5	AT3G21060	CPS50	
	ASH2L	AT1G51450	CPS40/ CPS60	

*For references see text.

the multicellular filamentous fungi (euascomycetes, *Pezizomycetes*). Systematic analysis of the *SET* domain genes across the entire phylum has outlined clear distinctions between *SET*-domain gene collections in the unicellular and their multicellular relatives. Elaboration of multicellularity in metazoa and in plants has been accompanied by further expansion of the numbers and types of *SET*-domain gene families. The majority of the families found in the filamentous fungi are related to the families found in extant animals and plants suggesting common ancestral origins (Veerappan *et al.,* 2008).

Unicellular yeasts carry histone marks, and genes establishing them, associated mainly with transcriptional activation (Garcia *et al.,* 2007; Morris *et al*., 2007). Thus, hemiascomycetes carry only "activating" (H3K4, H3K36 and H3K79) marks, while genes associated with silencing (*SET9, MYND- SET*, and *Su(var)3-9)* have been lost (Dujon *et al.,* 2004; Cliften *et al.,* 2006). A paradigm is the loss of the *Su(var)3-9* gene, and of the respective H3K9me mark, resulting in disappearance of the entire machinery making heterochromatin. This phenomenon raises the important question of how yeasts silence their genes and genome domains. Answers are suggested by the remarkable ability of yeasts to adopt available means to achieve ends that are functionally similar but molecularly different from mechanisms employed by other systems. For example, *S. cerevisiae* assembles silencing chromatin structure by a principally different molecular approach: through binding of transcription factors (RAP1), of silencers (Sir1, 2, 3, 4), and a component of the replication machinery (ABF1, ORC1) to specific DNA sequences, it effectively substitutes for the lost heterochromatin machinery (Rusche *et al.,* 2003). Even more surprising is that close relatives of *S. cerevisiae* do not use the same tools but have evolved species-tailored mechanisms for achieving effects functionally similar to heterochromatin: there is no Sir1 in *C. glabrata*, no Sir1 and Sir3 in *A. gossypii*; neither Sir nor the RNAi-pathways are conserved in *D. hansenii* and none of the *S. cerevisiae* heterochromatin factors was found in *Y. lipolytica* (Fabre *et al.,* 2005).

Another approach compensating for lost *SET* domain genes in yeasts is taking advantage of the degree (mono-, di-, or tri-) methylation of the lysine4-NH2- groups to achieve different transcriptional outcomes for pertinent genes (Bernstein *et al.,* 2002; Santos-Rosa *et al*., 2002). In *S. cerevisiae,* the amount of methyltags on the same lysine residue is a sufficient signal for recruitment of repressive or activating complexes, while the fission yeast, the filamentous fungi, and higher multicellular systems use additional signs, in a complex 'code' carried by a larger number and diversity of histone marks.

Collectively, available data point to a great evolutionary divergence of 'invented' mechanisms to produce silencing effects in systems that have lost silencing epigenetic marks.

The appearance of the E(Z) family and of histone H3K27me3 mark

Given the tightly balanced activities of the Polycomb group (PcG) and the Trithorax group (TrxG) complexes in maintaining gene expression patterns in development, it is interesting to analyze the evolution of Trithorax- and of EZ-related functions in parallel. Phylogenetic analysis revealed that *Enhancer-of-zeste, E(z)*, genes are not present in the unicellular yeasts but *E(z)*-related sequences exist in the filamentous fungi, although clustering with the animal and in plant sequences with a low bootstrap (Veerappan *et al.,* 2008). Furthermore, K27me3 marks were not found in the genomes of filamentous fungi, supporting the idea that methylation of H3K27 illustrates a chromatin mark, possibly, associated with the evolution of highly specialized functions.

The structure of the SET domain peptides of the E(z)-type, the loss of the postSET domains in particular, is consistent with the appearance of a novel substrate specificity (Zhang *et al*., 2003). One important difference between animal and plant E(z) proteins, on the one hand, and the putative E(z)-related fungal proteins, on the other hand, is the conservation of the peptide sequences upstream of SET in animal and plant E(z) proteins. These domains are not conserved in the fungal proteins providing a possible reason for the absence of K27me3 marks on the fungal histone H3. The nature of the domains has not been fully resolved but they participate in the assembly of the PRC2 complexes (Cao and Zhang, 2004). It is interesting to note that phylogenetic analysis clustered animal and plant E(z) proteins (with bootstrap values of 99%, Veerappan *et al.,* 2008) suggesting that the last shared *E(z)*-ancestor containing these domains has existed before the separation of the animal and plant kingdoms. Moreover, this common ancestor should have occurred after the separation from the filamentous fungi; alternatively, the primordial *E(z)*-related gene carrying the conserved upstream domains might have been lost in the fungal lineage. Absence of *E(z)*-related genes in unicellular yeasts indicates that Polycomb mechanisms do not operate in these organisms. However, abundant H3K27me marks were found in *Tetrahymena* (Garcia *et al.,* 2007) suggesting that these marks may be serving unicellular species-specific needs as well. It will be important to establish whether the *Tetrahymena* E(z) proteins assemble Polycomb-group complexes and whether they function similarly to their multicellular relatives.

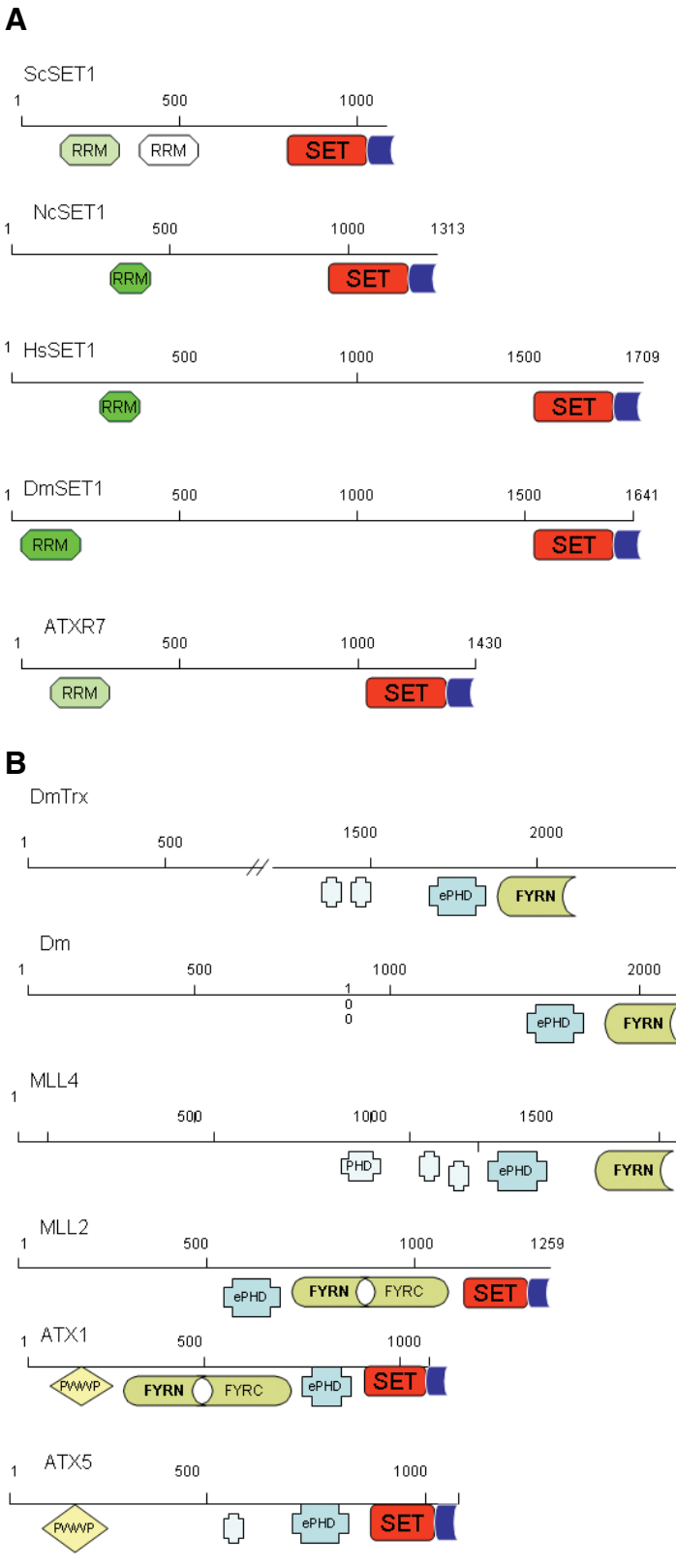
Gene duplication has provided opportunities for the evolution of multiple complexes with nuanced specialization of function. Three *Arabidopsis* E(Z)-like proteins, CURLY LEAF (CLF), MEDEA (MEA) and SWINGER (SWN) can assemble different PRC2-like complexes to regulate specific processes in the plant (Table 1; Chanvivattana *et al.,* 2004; Makarevich *et al.,* 2006; Schubert *et al.,* 2005; 2006). It is important to note also the multiplication of the EMF2/VRN2/FIS2 proteins specifically binding the CLF/SWN/MEA in each of the different PRC2 complexes (Table 1). The rapid evolution of *MEA* has suggested neofunctionalization after the duplication of the ancestral E(Z) homolog, supporting a later origin of genomic imprinting within *Brassicaceae* species (Spillane *et al.,* 2007).

Evolution of the Trithorax function in unicellular and multicellular organisms

The *Set1* gene of *S. cerevisiae* encodes a member of the large TRITHORAX family (named after the *Drosophila* Trithorax protein). The SET domain and the adjacent cysteine-rich motif (postSET) are the two most highly conserved sequences defining a protein's belonging to the family (Fig. 1). Phylogenetic analysis clearly distinguishes two subfamilies: the SET1 and the Trithorax. Importantly, in the genomes of both unicellular and filamentous fungi, the family is represented by a single copy of the SET1-type; by contrast, animal and plant genomes contain multiple copies of

the *SET1*-and of the *Trithorax*-subtypes.

The SET1 subfamily



The ancestral gene encoding SET1 in *S. cerevisiae* has been conserved throughout the evolution and is present in all unicellular and multicellular species. The SET-postSET sequences (located at the C-terminus) are highly conserved in all proteins (Fig. 1). The upstream regions are highly conserved in hemiascomycetes but are divergent in the SET1 proteins of *S. pombe* and the *Y. lipolytica*. These sequences encode an RRM (RNA-recognition motif) found also in the SET1 proteins of the filamentous fungi and in the *SET1* orthologs of animal and plant genomes. One copy of the *SET1*-subtype is found in *Drosophila*, two in mouse and humans (an apparent duplication), and one in *Arabidopsis* (the *ATXR7* gene, *At5g42400*). The animal and the plant SET1-counterparts are related to the fungal SET1 proteins and cluster within the same phylogenetic group (Veerappan *et al.,* 2008). Two genes in *Chlamydomonas*, *SET1* and *SET4,* are related to the yeast *SET1* but, apparently, do not carry RRM-encoding sequences (van Dijk *et al.,* 2005). These *SET1* and *SET4* were associated with generation of the K4me1 and K4me2/3 marks, respectively.

Thereby, in the genomes of unicellular organisms, of filamentous fungi, and of higher eukaryotes, the *SET1*-related genes are, most likely, orthologs involved in 'core' cellular activities not connected with functions required for multicellularity (Aravind and Subramanian, 1999). In a remarkable contrast, the genes from the Trithorax (TRX) subfamily are not represented in the genomes of unicellular and filamentous fungi. It is important to emphasize that members of the TRX subfamily, but not of the SET1, participate in the antagonistic TrxG/PcG complexes (Table 1).

The TRX subfamily

The members of this group carry SET-postSET regions highly related to the proteins from the SET1-subfamily. Most likely, the ancestral *SET1*-related gene has multiplied and diversified its structure (and, thereby, function) after separation from the lineages carrying only the *SET1* gene. Signature structural motifs of the TRX subfamily members are the PHD domains and the FYRN- and FYRC (DAST) domains (Alvarez-Venegas and Avramova, 2001; Fig. 1B). The two DAST motifs may be located adjacently, or spread apart. The roles of these motifs are largely unknown but acquisition of new

Fig. 1. Architecture of the SET1- and Trithorax (TRX) subfamilies across kingdoms. (A) *Members (paralogs) of the SET1-subfamily. Highly conserved sequences are color-coded. The less conserved RRM domain in yeast SET1 and in the* Arabidopsis *ortholog (ATXR7) are shown in paler colors. Domains are not drown to scale.* **(B)** *The Trithorax subfamily as found in animal and plant genomes. The two major sister groups differ by the positioning of the FYRN-FYRC domains (DAST). Two paralogous pairs found in the fly* (D.melanogaster, Dm)*, human (MLL4 and MLL2) and* Arabidopsis *(ATX1 and ATX5) genomes are shown. Notable is the presence of conserved domains, in addition to the SET-postSET regions, in the animal and the plant representatives.*

building blocks may reflect the evolution of SET1-related proteins in animals and plants to meet requirements for novel functions. Thereby, the ancestor of the animal and plant TRX-lineages might have occurred after the separation of the fungal SET1-branch, as discussed above for the ancestor of the animal and plant *E(z)* genes. Apparently, a primordial version of the antagonistic PcG/TrxG mechanism has existed in the common ancestor before the separation of the animal and plant kingdoms. However, whether the ancestral *E(z)/Trithorax* genes have appeared in an ancestor after the separation from the fungi, or whether these genes were lost at the branching off from the common ancestor, are fascinating questions that are remaining open.

Evolution of the *Trithorax* genes in Plants

Multiplication of an ancestral *TRX*-gene in *Arabidopsis* has produced five copies clustered in two sister groups: ATX1 and ATX2 forming one group, and ATX3, ATX4, and ATX5 forming the second (Baumbusch *et al.*, 2001; Alvarez-Venegas and Avramova, 2002). In rice, one protein, XP_450166 (SDG723) is a putative ortholog of both ATX1/ATX2, while the rice NP_913370 clusters with the ATX3/ATX4/ATX5 sister group (Fig. 2; Ng *et al.*, 2007). Apparently, the divergence of the two sister groups has taken place before the separation of the mono- and the di-cots. The respective maize homologs (Springer *et al.*, 2003) are available only as short peptides and could not be clustered with confidence.

A defining structural feature separating the two subgroups is the presence of DAST (the FYRN/FYRC juxtaposed version) in the ATX1/2 sister group; ATX3, ATX4, and ATX5 do not have DAST but carry an additional PHD finger. The function of the DAST peptides is not known but its presence/absence in the plant trithorax group underlies the segregation of the *Arabidopsis* proteins into two sister clades (Fig. 2).

A subset of animal Trithorax homologs, including insects, vertebrates, mammalian, carry juxtaposed DAST domains similar to the ATX1/ATX2 subgroup. However, split-DAST paralogs are present in the same genomes as well; for example, the *Drosophila* protein Trithorax, and the mammalian MLL1 and MLL2 belong to a sister group containing trithorax proteins with spread-apart DAST motifs (Fig. 1B).

In addition to the ATX family, seven *Arabidopsis* proteins have been classified as Trithorax-Related, ATXR (Baumbusch *et al.*, 2001). Our phylogenetic analysis, however, identified only ATXR7 as a Trithorax family member representing the *Arabidopsis* ortholog of SET1 (see above; Fig. 1A); the AAN01115 protein (encoded by the *Os12g41900* gene) is the SET1 ortholog in rice. The other ATXR proteins cluster in separate groups distantly related to Trithorax (Fig. 2). In contrast to an earlier report that ATXR5/ATXR6 belong in the SET3/SET4 group of *S. cerevisiae* (Springer *et al.*, 2003) our phylogenetic analysis failed to cluster them together. Furthermore, detailed comparative analyses revealed that the ATXR5/6-SET domain sequences do not carry the hallmark amino acid substitutions defining the SET3 subfamily (Veerappan *et al.*, 2008). Enzyme activity has not been established but ATXR5 and ATXR6 differ in subcellular localization and functions. The two paralogs interact with proliferating cellular nuclear antigen (PCNA) and are critically involved in DNA replication, DNA repair, maintenance, and heterochromatin formation (Raynaud *et al.,* 2006).

Proliferation of the *ATX* genes illustrates an evolutionary trend implying exclusive roles for these family members in *Arabidopsis*. Gene duplication, followed by functional divergence of the resulting pair of paralogous proteins, is a major force shaping molecular networks in living organisms. Duplicated genes involved in transcriptional regulation might have been preferentially retained leading to the origination of a non-overlapping pathway to function in two different cell types, developmental stages, or environmental conditions (Blank and Wolfe, 2004). Epigenetic regulators can modulate expression of a large number of functionally linked genes suggesting that a duplicated epigenetic factor might be critically linked with the evolution of novel regulatory networks. Duplicated genes (paralogs) may acquire different fates including silencing or null mutations, partitioning of functions to complement the range of activity of the ancestral gene, or gain of functions. The two paralogs might have parceled out the range of

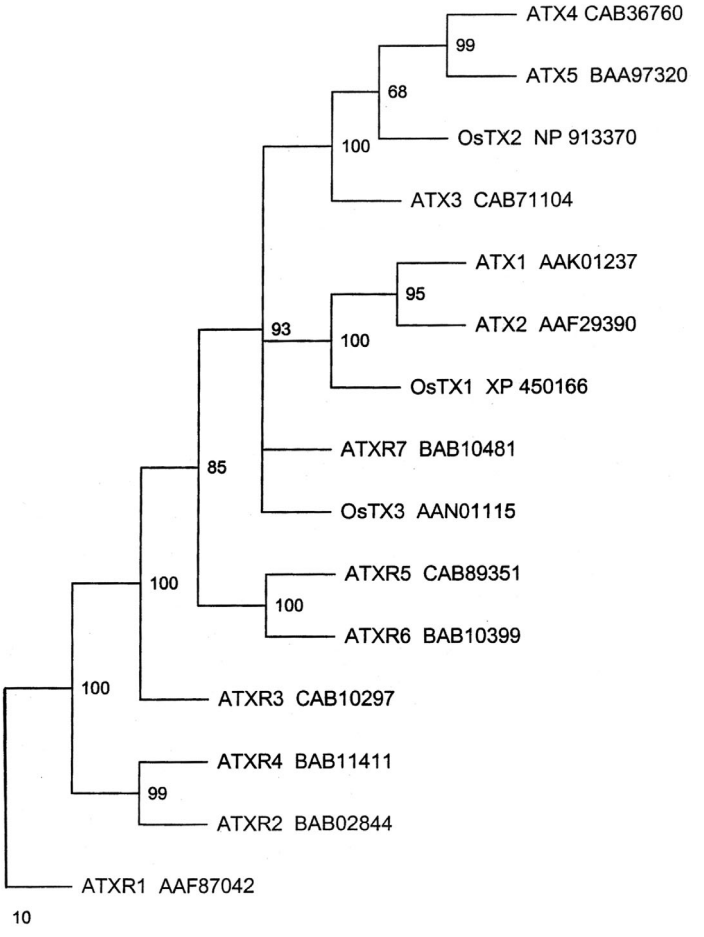


Fig. 2. Maximum parsimony (MP) phylogeny of the Trithorax (TRX) family within *Arabidopsis* and rice genomes. *Sequences of the proteins from the* Arabidopsis *and rice TRX family were used to reconstruct MP trees. Figures illustrate bootstrap values (500 pseudoreplicas;100=100%). Only bootstrap values greater than 60% are shown. All proteins from the* Arabidopsis *Trithorax-Related (TRXR) family, except ATXR7, cluster separately from the ATX family showing a more distant relationship. ATXR7 and AAN01115 are the SET1 orthologs in* Arabidopsis *and rice, respectively. Proteins are identified by their given names and respective accession numbers.*

pleiotropic functions of the ancestral gene so that each copy now contributes a portion of ancestral activities. The latter path may lead to full separation of functions, or neofunctionalization (Kondrashov *et al.*, 2002). While it is logical to expect that paralogs encoding divergent proteins have evolved novel functions, it is impossible to predict the outcomes from duplicated genes with highly preserved coding sequences. For this reason, it was particularly revealing to compare the fates of the two most closely related genes, *ATX1* and *ATX2*.

Origin of* ATX1 *and* ATX2 *genes

The *ATX1* and *ATX2* paralogs have originated from a segmental chromosomal duplication (Baumbusch *et al.,* 2001) after the separation of the dicots from the monocots (Ng *et al.*, 2007). ATX1 and ATX2 are 65% identical, 75% similar at the amino acid level. The two proteins are built by similar architectural motifs. However, homozygous mutant *atx1* and *atx2*-mutant lines display strikingly different phenotypes: in contrast to the early bolting and numerous flower-organ aberrations of *atx1* plants (Alvarez-Venegas *et al.,* 2003), mutant *atx2* lines did not differ detectably from the wild type (Saleh *et al.*, 2008b). At a first glance, the inability of ATX2 to substitute for ATX1 in the *atx1* background might suggest nonfunctionalization following the duplication: extant *ATX1* continues to play the ancestral function, while ATX2 has become non-essential. Further analyses, however, revealed complex relationships between the ATX1 and ATX2 paralogs.

***Promotor divergence of* ATX1 *and* ATX2**

By being expressed uniquely in different temporal and/or spatial manner, redundant genes may acquire functional independence. Changes in *cis*-regulatory elements, then, might reflect steps toward functional divergence (Kondrashov *et al.*, 2002). In this context, the larger number (~17) of recognized transcription factor (TF) binding 'motifs' in the *ATX1* promotor, compared to six at the promotor of *ATX2,* suggests that *ATX1* expression would be controlled by a potentially larger set of TFs than *ATX2.* Consistent with a divergent set of regulatory elements driving their expression in temporally and spatially different domains, *ATX1* and *ATX2* display non-overlapping or partially overlapping expression domains (Saleh *et al.*, 2008b). In particular, double *atx1/atx2* mutants in the *FRI* background suppressed the late flowering phenotype more dramatically than the single atx1 mutant suggesting that ATX1 and ATX2 play a partially redundant role in activating *FLC* (Pien *et al.*, 2008). Thereby, selection has preserved the two structurally similar proteins with diverged *cis*-regulatory sequences. Similar regulation is possible as well, as both promotors carry putative sites for regulation by light, UV radiation, pathogen attacks, wounding, abscisic acid, etc.

Functional divergence of ATX1 and ATX2; non-redundant roles of ATX1 and ATX2 in overall gene control

Regulation of homeotic genes is just one possible role for Trithorax proteins. Estimated by the conservative Bonferroni method restricting False Discovery Rate, 867genes (~7% of all detectably expressed genes) changed expression in the *atx1* background, consistent with the pleiotropic role of ATX1 (Alvarez-Venegas *et al.,* 2006a); about 0.7% (80 genes) changed expression in the *atx2* background. These results provide evidence that *ATX2* has remained functional but with a more restricted role in *Arabidopsis. ATX1*- and *ATX2*- regulated genes cover a broad spectrum of similar functions. Despite this apparent redundancy, however, cluster analysis revealed that ~60% of the ATX2-targets were not shared with ATX1 representing, thus, ATX2-specific targets. Within the shared 34 gene-set, 26 genes changed expression in opposite directions; only 8 genes were co-regulated (1 up-, 7 down-regulated). Further analysis revealed that even within the shared set, ATX1 and ATX2 employ different mechanisms for their regulation (Saleh *et al.*, 2008a,b).

Evolution of the H3K4 methyl marks in plants

In yeasts, SET1 is the sole activity responsible for the overall chromatin modification and for establishing mono-, di-, and trimethyl-H3K4 marks (Bernstein *et al.*, 2002; Santos-Rosa *et al.*, 2002). In *Chlamydomonas*, a *SET1*-encoded activity deposits a K4-monomethyl mark (van Dijk *et al.*, 2005). SET1-orthologs are present as single copies in the genomes of the filamentous fungi suggesting that these organisms use H3K4me-mechanisms similar to those of yeasts. By contrast, known animal and plant Trithorax enzymes modify only a limited fraction of target nucleosomes (Alvarez-Venegas and Avramova, 2005; Wysocka *et al.*, 2005; Ruthenberg *et al.*, 2007) implying involvement of multiple K4 methyltransferases. The degree of methylated K4-NH2 groups may serve different roles: di- or tri-methylated H3K4 are marks for non-active, or active genes, respectively, in *S.cerevisiae* (Bernstein *et al.*, 2002; Santos-Rosa *et al.*, 2002), while in metazoa, both modifications label actively transcribed genes (Santos-Rosa *et al.*, 2002; Kouzarides, 2007). In *Arabidopsis*, H3K4me2 marks are found at coding regions independent of whether the gene is active or not; their absence from intergenic regions suggested that H3K4me2 could be a general mark distinguishing transcribed from non-transcribed sequences in the genome (Alvarez-Venegas and Avramova 2005). In maize, H3K4me2 is limited to areas demarcating the euchromatin gene-space (Shi and Dawe, 2006). Despite the broad distribution of H3K4me2 marks in *Arabidopsis* euchromatin (Jasencakova *et al.*, 2003), no activity involved in di-methylating H3K4 marks has been identified.

ATX2 was involved in di-methylating H3K4 at selected loci providing the first example of a K4 di-methylase separate from the K4 tri- methylating activity (Saleh *et al.*, 2008b). In rice, there is one gene in the ATX1-ATX2 sister clade, suggesting that in the monocot, one gene might be playing the roles of two genes in *Arabidopsis*. Initially redundant, ATX1 and ATX2 may have evolved through splitting the functions of an ancestral gene consistent with neofunctionalization models (Kondrashov *et al.*, 2002).

H3K4-trimethyltransferase activity was reported for the *Arabidopsis* protein, EFS (Kim *et al.*, 2005). Another group, however, found that this protein displayed specificity for H3K36, but not for H3K4 (Zhao *et al.,* 2005). It is worth noting that the SET domain of EFS is of a different ancestry that the Trithorax lineage, belonging to the ASH family (Baumbusch *et al.*, 2001).

The specific 'language' of the histone marks

Different biological systems have evolved different ways of implementing the histone marks suggesting that the 'language' is species-specific (Loidl, 2004). For instance, H4K20me in metazoa may provoke a series of events leading to the formation of heterochromatin (Schotta *et al.*, 2004), but in *S.pombe,* the mark

is a signal for DNA repair (Du *et al*., 2006); hemiascomycetes and plants (Zhang *et al.,* 2007b) are void of this mark altogether, and *SET9*/*Su(var)4-20* genes were not found in *Arabidopsis*, rice, and maize genomes. Furthermore, there is no *Polycomb* homolog in the *Arabidopsis* genome and a different mechanism 'reads' the H3K27 modifications (Schubert *et al*., 2006; Turck *et al*., 2007). In maize, H3K9me3 is associated with centromeres, while H3K9me2 and H3K27me3 occur in euchromatic domains, but H3K9me2 does not overlap with either H3K27me3 or H3K4me2 (Shi and Dawe, 2006). Plants have a unique family of chromodomain DNA-methyltransferases, a unique HDAC family and differentially modified histone lysine residues than other known chromatins (Loidl, 2004 and ref. therein; Zhang *et al.,* 2007b).

Despite the specific usage of the histone-tail marks, acetylated histones and methylated histone H3 lysines 4 and 36 are generally associated with transcribed genes, while deacetylated histones and methylated lysines 9 and 27 are representing silent loci (Kouzarides, 2007). Amounting new evidence, however, is pointing to correlations more complex than simply 'activating/silencing' tags. For example, deacetylation of the coding regions in transcribed genes was linked directly with active transcription and with histone H3K36me2, a mark of actively transcribed genes (Keogh *et al*., 2005); simultaneously present H3K4me3 and H3K27me3 marks found at silent genes in embryonic stem cells suggested that co-existing 'activating' and 'silencing' nucleosomal modifications establish a bivalent chromatin state at loci 'poised' for transcription later in development (Bernstein *et al*., 2006). It is remarkable that the chromatin at the flower homeotic gene locus, *AG*, is similarly tagged by H3K4me3 and H3K27me3 in its silent state (Saleh *et al*., 2007), suggesting that dual methylations might be chromatin marks for genes involved in plant developmental processes as well. Furthermore, *Arabidopsis* genes may carry methylated H3K9, H3K27 and H3K4 in various combinations in a gene-, tissue- or developmentally controlled patterns. Absence of H3K4me3 tags does not necessarily correlate with low expression levels (Alvarez-Venegas and Avramova, 2005; Saleh *et al*., 2008b). Thereby, correlations between histone methylation profiles and gene activity appear to be much more complex than initially perceived. Whether histone H3-tail lysine methylation modifications precede or trail established transcriptionaly active states remains to be seen.

Involvement of Trithorax in development, disease response and cell signaling in *Arabidopsis*

Similar to Trithorax activity in animals systems, *ATX1* regulates plant homeotic genes controlling flower organ formation and organ identity. The involvement of *ATX* in disease-response mechanisms, in cell signaling, and in regulating the transition from vegetative to flowering stages illustrates the pleiotropic roles played by the plant counterpart of trithorax.

Antagonistic PcG/TrxG functions in Arabidopsis

The *Arabidopsis* homolog of E(z), CLF, suppresses the expression of the flower homeotic gene *AG.* Derepression and ectopic expression of *AG* in leaves is partly responsible for the curly leaf and early flowering phenotypes (Goodrich *et al*., 1997). In contrast, ATX1 upregulates *AG;* lower *AG* transcript levels are associated with multiple organ malformations and earlier bolting of *atx1* plants (Alvarez-Venegas *et al*., 2003; Saleh *et al.,* 2007). Introduction of *atx1*$^{-/-}$ in the *clf* background, however, shifted the phenotypes of homozygous double mutant plants towards the wild type. Thus, loss of both ATX1 and CLF functions rescued the single-mutant phenotypes (Saleh *et al.,* 2007) providing evidence for coordinate antagonistic participation of plant PcG/TrxG factors in the regulation of a shared gene locus. ATX1 and CLF physically interact providing mechanistic basis for the observed effects. Partial normalization of axial–skeletal transformations in mice was also observed when *Mll* (a human homolog of trithorax) and *BMI*-1 (a PcG component) were simultaneously deleted (Xia *et al*., 2003).

At the molecular level, both H3K4me3 and H3K27me3 marks were required for the normal suppression of *AG* and for establishing a chromatin state similar to the bivalent states of embryonic stem cell chromatin (Bernstein *et al*., 2007). Contrary to the expectation that absent *ATX1* and *CLF* functions would erase the H3K4me3 and H3K27me3 marks, there was a partial restoration of the marks on the *AG*-nucleosomes in the double-mutant chromatin. These results suggested that in *Arabidopsis,* PcG and TrxG complexes function in specific pairs (Saleh *et al.,* 2007). Given that a functionally equivalent methylase cannot substitute a missing relative within a complex, the model takes the specificity one step further by suggesting that a PcG - and a TrxG-complexes function as a specific pair to establish bivalent marks at particular loci. Only when both complexes were missing could their roles be undertaken by another pair. However, restored modifications were not identical with the wild type nor did they take place 100% of the times. These observations offer clues for interpreting spontaneous reversals, variability and instability of phenotypes associated with epigenetic mutations.

ATX1 in organ primordia development and organ identity

ATX1 activity is required to maintain normal levels of homeotic gene expression during flower development (Alvarez-Venegas *et al*., 2003). Homeotic transformations showing stamenoid petals and carpeloid stamens in *atx1* mutant flowers resulted from conversions of second-whorl into third-whorl organs, and of third-whorl into fourth-whorl structures, respectively (Alvarez-Venegas *et al*., 2003). The pleiotropy and variable expressivity of the *atx1* phenotype makes it difficult to assign a developmental defect to a particular compromised homeotic function. The phenotypes resulting from the *ATX1* loss-of-function were relatively mild, if compared to the drastic effects of homeotic mutants themselves. This might be related to the fact that *ATX1* regulates the expression of homeotic genes with counteracting functions: for example, lowered expression of a class C- gene could counterbalance effects resulting from down-regulated class A and class B functions. Alternatively, *atx1* does not affect homeotic gene expression uniformly but showed temporal and spatial differences, possibly accounting for the weaker phenotype as compared to homeotic mutations.

Transition to flowering

The *FLC* gene is positioned at the convergence nod of at least four distinct pathways blocking transition from vegetative to reproductive stage in *Arabidopsis* (He and Amasino, 2005 and ref. therein). Nucleosomes of the silent *FLC* carry deacetylated histones (He and Amasino, 2005 and ref. therein), di-methylated

H3K9, di- and tri methylated H3K27 (Bastow *et al*., 2004; Chanvivattana *et al*., 2004; Schubert *et al*., 2005). The SET domain containing protein, AtCZS, is involved in establishing the HK27me2 marks (Krichevsky *et al.,* 2007)**,** while a PRC2-like complex containing the E(z)-homologs CLF or SWN suppresses *FLC* expression (Chanvivattana *et al*., 2004; Schubert *et al*., 2005; 2006) even in nonvernalized plants (Wood *et al.,* 2006). On the other hand, nucleosomes of transcribed *FLC* carry di-methylated H3K4 (Bastow *et al*., 2004), tri-methylated K4 (Kim *et al*., 2005), methylated K36 (Zhao *et al.,* 2005) and symmetrically dimethylated H4R3 (Wang *et al*., 2007). Suppressed demethylation of K4me2 paralleled an increase in *FLC* transcripts supporting a link between lysine 4-methylation and transcriptional competence at this locus (Jiang *et al.,* 2007; Liu *et al*., 2007). A chromatin-based mechanism involving the histone variant, H2A.Z (Deal *et al*., 2007; March-Díaz *et al.,* 2007) controls *FLC* transcription illustrating the involvement of chromatin structure and histone modifications in regulating the transition to flowering.

Collectively, the data underscore a correlation between *FLC* expression state and the presence of 'activating/silencing' marks on *FLC* nucleosomes. It was surprising, then, to find coexisting H3K4me3 and H3K27me3 marks distributed throughout the *FLC* coding sequence (Saleh *et al.,* 2008a). Apparently, the dual marks did not impede transcription, consistent with a model that Trithorax may function as an anti repressor of PcG preventing inappropriate silencing of *HOX* genes (Papp and Mueller, 2006). Furthermore, the methylation pattern at the 5'-end transcription start nucleosomes changed dynamically with changes in transcriptional activity, while the pattern on downstream gene nucleosomes remained stable throughout developmental transitions. Similar patterns were displayed at the locus of a related MADS-box transcription factor, the *AP1* locus, as well. However, the dynamic changes at the *AP1* 5´-end transcription start site, involved removal of a nucleosome in a developmentally regulated process, while downstream regions remained stably labeled by both K4me3 and K27me3. ATX1 is directly involved in 'writing' the H3K4me3 marks on'*FLC*, but not on *AP1,* nucleosomes indicating that its effect on *AP1* is indirect (Saleh *et al.,* 2008a). CLF does not seem involved in modifying either *FLC* or *AP* leaving open the question of which EZ activity trimethylates H3K27me3 on *FLC* and *AP* nucleosomes.

Epigenetic regulation of disease-response mechanisms in Plants

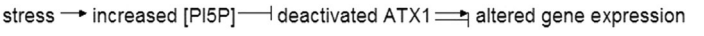
The nucleosome-remodeling factor DDM1 *(diminishes DNA methylation)* and DNA methylation are involved in plant-pathogen interactions controlled by the *BAL* locus (Stokes *et al*., 2002). Histone deacetylases were implicated in the JA-mediated defense-responses of *Arabidopsis* (Devoto *et al*., 2002), and histone H3 methylation was implicated in the plants' response to *Pseudomonas syringae* (Alvarez-Venegas *et al*., 2007a). The *WRKY70* gene encoding a transcription factor at the cross section of the antagonistic SA- and JA-signaling pathways (Li *et al*., 2004) is a direct target of the histone modifying activity of ATX1; the downstream SA-responsive gene, *PR1,* and the JA-responsive gene *THI2.1* are indirect (secondary) targets (Alvarez-Venegas *et al*., 2007a). Importantly, non-induced *PR1* and *THI2.1* genes displayed H3K4me3-modified nucleosomes. This observation suggested that defense-response genes might keep their nucleosomes in actively modified 'ready' state for a quick-change in transcription upon need. This could provide a molecular mechanism for rapid coordinated changes in expression of entire gene networks. Nucleosomal tags, however, are not sufficient to initiate transcription on their own: regulation by TFs is superior to the methylation profiles in initiating gene activation or repression (Alvarez-Venegas *et al.,* 2007a).

ATX1 links epigenetic regulation with lipid signaling

ATX1 can bind specifically the lipid messenger phosphoinositide 5-phosphate, PtdIns5P, *via* its PHD finger (Alvarez-Venegas *et al.,* 2006a; 2006b). Phosphoinositide phosphates (PtdInsP) are important components of the cell lipid pool acting as intracellular and intercellular messengers in processes mediating plant growth, development, cytoskeletal rearrangements, and signal transduction (Wang, 2004). Existence of diverse phosphorylated isomers creates selective means for communication and for coordinating cell growth (Stevenson *et al*., 2000). The monophosphorylated isoform, PtdIns5P, is a distinct minor component of the cellular phosphoinositide pool increasing its levels in response to hyperosmotic stress (Meijer *et al.,* 2001). It may also serve as a precursor for phosphatidyl inositol-biphosphates, PtdIns3,5P_2 and PtdIns4,5P_2, the synthesis of which also increases rapidly when cells respond to hyperosmotic stress (Meijer *et al.,* 2001).

The highly conserved PHD-peptide is present in many nuclear proteins belonging in distinct families with different functions: the PHD domain of the putative tumor suppressor (ING2) binds PtdIns5P, and to a lesser extent PtdIns3P, while the PHD of the repressor Mi2 did not bind any of the tested lipids (Gozani *et al*., 2003). The PHD fingers of ING2 and NURF bind H3K4me3 (Li *et al*., 2006; Wysocka *et al*., 2006). The PHD motifs of ATX1 and ATX2 belong in a subgroup defined as extended PHD, ePHD, conserved in all Trithorax members of animal and plant origin.

Binding of ATX1-PHD to PtdIns5P is responsible for the intracellular location of ATX1 suggesting that its nuclear localization may depend upon factors affecting the concentration of PtdIns(5)P. The latter might be environmentally induced (Meijer *et al.,* 2001) and might be under cell cycle or developmental control (Clarke *et al.,* 2001). A distinct set of ATX1 and PtdIns5P co-regulated genes was identified by whole-genome expression profiling (Alvarez-Venegas *et al*., 2006a) providing biological relevance for their interaction. Thereby, ATX1 is not a constitutively nuclear protein and its subcellular localization might be a tissue or cell-specific event reflecting changes in response to internal and/or external signals. Changes in PtdIns5P levels might have major implications for the activity of ATX1 and for the expression of the shared target genes. A model (summarized below) suggests a pathway that translates stress into altered gene expression. It involves a signal (PtdIns5P) and a receptor (ATX1) modifying the expression rates of targeted response-related genes (Alvarez-Venegas *et al*., 2006a).



Lipid signaling is involved in functions other than a role restricted to the plasma membrane (Jones *et al.,* 2006; Gozani

et al., 2003). PtdIns5P negatively affects ATX1 function linking epigenetic regulation with lipid signaling.

Perspectives

PcG/TrxG complexes are involved in plant developmental and adaptation processes. Current studies of plant epigenetic regulation, however, have been focused mainly on the gene-silencing component of the mechanism. A deliberate shift in focus towards gene-activating mechanisms would be required before we begin to comprehend epigenetic regulation in plants. Given the low activity of recombinant Trithorax proteins, to study biochemically Trithorax functions in plants would require isolation of ATX binding partners and of a Trithorax complex similar to the animal and the yeast COMPASS counterparts (Table 1; Steward *et al*., 2006, and ref. therein). The finding that ATX1 and CLF interact at the *AG* locus illustrated an unknown ability of a TRX homolog to bind directly an E(z) homolog and suggested that other PcG/TrxG specific pairs are likely to be recognized in the future. The presence of dual activating/silencing marks (Saleh *et al*., 2007), as well as the reported crosstalk between histone H2B ubiquitylation and H3K4me3 in *Arabidopsis* (Sridhar *et al*., 2007) illustrated features of the 'code' conserved in animal and plant systems. However, given the specific modifications of plant histones (Zhang *et al*., 2007), uncovering correlations between histone modifications and plant-specific 'dialects' would continue to be an exciting research.

While it is clear that chromatin structure influences gene activity, understanding how histone modifications and chromatin remodeling are related mechanistically and how epigenetic regulators achieve control over a large number of gene targets are among the most complex matters. Recent studies have provided clues as to how chemical modifications of histones might control expression at the mechanistic level. Chromatin remodelers (SWI/SNF, RSC, ISWI, INO80, SWR1 and Mi-2/CHD groups), remarkably conserved throughout eukaryotes, are specialized multiprotein machines enabling access to nucleosomal DNA by altering the structure, composition, and positioning of nucleosomes (rev. in Saha *et al*., 2006). ATP-dependent activities can remodel chromatin by either mobilizing nucleosomes on DNA or by exchanging one histone variant for another, within the nucleosome (Reyes, 2005 and ref. therein). Nucleosome shifting and remodeling mechanisms are, apparently, linked. For example, the four-subunit NURF complex, critically involved in development, facilitates the formation of the preinitiation complex through the remodeling of nucleosomes at the promotor. BPTF recognizes K4me2/3 *via* the PHD domain of its largest subunit (Wysocka *et al*., 2006) providing a direct link for the activating role of K4me at the mechanistic level. It remains to be seen whether similar mechanisms function in plant chromatin.

An epigenetic regulator may influence the expression of specific members even within the same family, posing a logistic problem of how the simultaneous, selective, and finely tuned control of a multitude of genes could be achieved. A plausible scenario is that most downstream components of a network have their nucleosomes appropriately modified residing in a state of readiness for a quick response. Ultimately, their expression state would be determined by the activity of a transcription factor (activator or inhibitor), which would be a preferred target for epigenetic remodeling. Thereby, epigenetic regulation may be viewed as superimposed on primary regulatory systems, achieved by transcription factors. Such a model provides flexibility and a means for a rapid change in transcripts from many genes without the need to modify individually each component (Alvarez-Venegas *et al*., 2007a).

Recently, it was suggested that epigenetic systems "will not initiate a change of state at a particular locus but would register a change already imposed by other events" (Bird, 2007). Further studies would be required to provide evidence for the validity of these models and for providing an answer to the compelling question of whether epigenetic marks are responsive or proactive.

Acknowledgement

I am grateful to Dr. M. Sadder (University of Amman, Jordan) for reconstructing the plant TRX-MP tree. Studies reported in this paper were partially supported by NSF-EPS-0701892.

References

ALVAREZ-VENEGAS, R., AL-ABDALLAT, A., GUO, M., ALFANO, J.P. and AVRAMOVA, Z. (2007a). Epigenetic Control of a Transcription Factor at the Cross Section of Two Antagonistic Pathways. *Epigenetics* 3: 106-113.

ALVAREZ-VENEGAS, R. and AVRAMOVA, Z. (2001). Two *Arabidopsis* homologs of the *Drosophila Trithorax* genes; Identification of a new structural domain. *Gene* 271: 215-221.

ALVAREZ-VENEGAS, R. and AVRAMOVA, Z. (2002). SET-domain proteins of the Su(var), E(z) and T*rithorax* families. *Gene* 285: 25-37.

ALVAREZ-VENEGAS, R. and AVRAMOVA, Z. (2005). Methylation Patterns of Histone H3 Lys 4, Lys 9 and Lys 27 in transcriptionally active and inactive *Arabidopsis* genes and in *atx1* mutants. *Nucl Acids Res* 33:5199-5207.

ALVAREZ-VENEGAS, R., PIEN, S., SADDER, M., WITMER, X., GROSSNIKLAUS, U. and AVRAMOVA, Z. (2003). ATX-1, an Arabidopsis homolog of trithorax, activates flower homeotic genes. *Curr. Biol.* 13:627-637.

ALVAREZ-VENEGAS, R., SADDER, M., HLAVACKA, A., BALUSKA, F., XIA, Y., LU, G., FIRSOV, A., SARATH, G., MORIYAMA, H., DUBROVSKY, J. and AVRAMOVA, Z. (2006a). The Arabidopsis Homolog of Trithorax, ATX1, Binds Phosphoinositide 5-Phosphate and the Two Regulate a Common Set of Target Genes. *Proc Natl Acad Sci USA* 103: 6049-6054.

ALVAREZ-VENEGAS, R., TIKHONOV, A., SADDER, M. and AVRAMOVA, Z. (2007b) Origin of the Bacterial SET Domain Genes: Vertical or Horizontal? *Mol Biol Evol* 24: 482-497.

ALVAREZ-VENEGAS, R., XIA, Y., LU, G. and AVRAMOVA, Z. (2006b). Phosphoinositide 5-Phosphate and Phosphoinositide 4-Phosphate Trigger Distinct Specific Responses of *Arabidopsis* Genes. *Plant Signaling and Behavior* 1:140-151.

ARAVIND, L. and SUBRAMANIAN, G. (1999). Origin of multicellular eukaryotes - insights from proteome comparisons. *Curr Opin Genet Dev* 9: 688-694.

AVRAMOVA, Z. (2002). Heterochromatin in animals and plants; Similarities and differences. *Plant Physiol* 129: 40-49.

BASTOW, R., MYLNE, J.S., LISTER, C., LIPPMAN, Z., MARTIENSSEN, R.A. and DEAN, C. (2004). Vernalization requires epigenetic silencing of FLC by histone methylation. *Nature* 427:164-167.

BAUMBUSCH, L.O., THORSTENSEN, T., KRAUSS, V., FISCHER, A., NAUMANN, K., ASSALKHOU, R., SCHULTZ, I., REUTER, G. and AALEN, R.B. (2001). The Arabidopsis thaliana genome contains at least 29 active genes encoding SET-domain proteins that can be assigned to four evolutionary conserved classes. *Nucleic Acid Res* 29: 4319-4333.

BERNSTEIN, B.E., HUMPHREY, E.L., ERLICH, R.L., SCHNEIDER, R., BOUMAN, P., LIU, J.S., KOUZARIDES, T. and SCHREIBER, S.L. (2002). Methylation of histone H3 Lys 4 in coding regions of active genes. *Proc Natl Acad Sci USA* 99:8695-8700.

BERNSTEIN, B.E., MIKKELSEN, T.S., XIE, X., KAMAL, M., HUEBERT, D.J.,

CUFF, J., FRY, B., MEISSNER, A., WERNIG, M., PLATH, K., JAENISCH, R., WAGSCHAL, A., FEIL, R., SCHREIBER, S.L. and LANDER, E.S. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125: 315-326.

BIRD, A. (2007). Perceptions of epigenetics. *Nature* 447: 396-398.

BLANC, G. and WOLFE, K.H. (2004). Functional divergence of duplicated genes formed by polyploidy during Arabidopsis evolution. *Plant Cell* 16:1679-1691.

BRINK, R.A. (1956). A genetic change associated with the *R* locus in maize which is directed and potentially reversible. *Genetics* 41: 872–889.

CAO., R. and ZHANG, Y. (2004). The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. *Curr Opin Genet Dev* 14: 155-164.

CATCHESIDE, D.G. (1938). A position effect in *Oenothera*. *J Genet* 38: 345-352; ibid. 48: 31-42.

CHANDLER, V.L. and STAM, M. (2004). Chromatin conversations: mechanisms and implications of paramutation. *Nat Rev Genet* 5: 532-544.

CHANVIVATTANA., Y., BISHOP., A., SCHUBERT., D., STOCK., C., MOON., Y.H., SUNG, Z.R. and GOODRICH, J. (2004). Interaction of Polycomb-group proteins controlling flowering in Arabidopsis. *Development* 131: 5263-5276.

CLARKE, J.H., LETCHER, A., D'ANTOS, C.S., HALSTED, J.R., IRVINE, R. and DIVECHA, N. (2001). Inositol lipids are regulated during cell cycle profession in the nuclei of murine erythroleukemia cells. *Biochem J* 357: 905-910.

CLIFTEN, P.F., FULTON, R.S., WILSON, R.K. and JOHNSON, M. (2006). After the duplication: gene loss and adaptation in *Saccharomyces* genomes. *Genetics* 172: 863-872.

DEAL, R.B., TOPP, C.N., MCKINNEY, E.C. and MEAGHER, R.B. (2007). Repression of flowering in Arabidopsis requires activation of FLOWERING LOCUS C expression by the histone variant H2A.Z. *Plant Cel.* 19: 74-83.

DEVOTO, A., ELLIS, C., MAGUSIN, A., CHANG, H.S., CHILCOTT, C., ZHU, T. and TURNER, J.G. (2002). Expression profiling reveals COI1 to be a key regulator of genes involved in wound- and methyl jasmonate-induced secondary metabolism, defense, and hormone interactions. *Plant Mol Biol* 58: 497-513.

DU, L-L., NAKAMURA, T.M. and RUSSELL, P. (2006). Histone modification-dependent and -independent pathways for recruitment of checkpoint protein Crb2 to double-strand breaks. *Genes Dev* 20: 1583 – 1596.

DUJON, B., SHERMAN, D., FISCHER, G., DURRENS, P., CASAREGOLA, S., LAFONTAINE, I., DE, MONTIGNY, J., MARCK, C., NEUVÃEGLISE, C., TALLA, E. *et al.,*57 co-authors) (2004). Genome evolution in yeasts. *Nature* 430: 35-44.

FABRE, E., MULLER, H., THERIZOLS, P., LAFONTAINE, I., DUJON, B. and FAIRHEAD, C. (2005). Comparative genomics in hemiascomycete yeast: evolution of sex, silencing, and subtelomeres. *Mol Biol Evo* 22: 856-873.

GARCIA, B.A., HAKE, S.B., DIAZ, R.L., KAUER, M., MORRIS, S.A., RECHT, J., SHABANOWITZ, J., MISHRA, N., STRAHL, B.D., ALLIS, C.D. and HUNT, D.F. (2007). Organismal differences in post-translational modifications in histones H3 and H4. *J Biol Chem* 282:7641-7655.

GOODRICH, J., PUANGSOMLEE, P., MARTIN, M., LONG, D., MEYEROWITZ, E.M. and COUPLAND, G. (1997). A Polycomb-group gene regulates homeotic gene expression in *Arabidopsis*. *Nature* 386: 44–51.

GOZANI, O., KARUMAN, P., JONES, D.R., IVANOV, D., CHA, J., LUGOVSKOY, A.A., BAIRD, C.L., ZHU, H., FIELD, S.J., LESSNICK, S.L., VILLASENOR, J., MEHROTRA, B., CHEN, J., RAO, V.R., BRUGGE, J.S. *et al.,*23 coauthors) (2003). The PHD finger of the chromatin-associated protein ING2 functions as a nuclear phosphoinositide receptor. *Cell* 114: 99-111.

GRANT-DOWNTON, R.T. and DICKINSON, H.G. (2005). Epigenetics and its implications for plant biology. 1. The epigenetic network in plants. *Ann Bot (Lond)* 96:1143-1164. ibid. 97:11-27.

GRIMAUD, C., NEGRE, N. and CAVALLI, G. (2006). From genetics to epigenetics: the tale of Polycomb group and trithorax group genes. *Chromosome Res* 14: 363-375.

HE, Y. and AMASINO, R.M. (2005). Role of chromatin modification in flowering-time control. *Trends Plant Sci* 10: 30-35.

HEDGES, S.B., BLAIR, J.E., VENTURI, M.L. and SHOE, J.L. (2004). A molecular timescale of eukaryote evolution and the rise of complex multicellular life. *BMC Evo Biology* 4:2 1-9.

HENDERSON, I.R. and JACOBSEN, S.E. (2007). Epigenetic inheritance in plants. *Nature* 447: 418-24.

JASENCAKOVA, Z., SOPPE, WJ.J., MEISTER, A., GERNAND, D., TURNER, B.M. and SCHUBERT, I. (2003). Histone modifications in Arabidopsis-high methylation of H3 lysine9 is dispensable for constitutive heterochromatin. *Plant J* 33: 471-480.

JONES, D.R., BULTSMA, Y., KEUNE, W.J., HALSTEAD, J.R., ELOUARRAT, D., MOHAMMED, S., HECK, A.J., D'SANTOS, C.S. and DIVECHA, N. (2006). Nuclear PtdIns5P as a transducer of stress signaling: an *in vivo* role for PIP4Kbeta. *Moll Cell* 23:685-695.

KAEPPLER, S.M., KAEPPLER, H.F. and RHEE, Y. (2000). Epigenetic aspects of somaclonal variation in plants. *Plant Mol Biol* 43:179-188.

KEOGH, M.C., KURDISTANI, S.K., MORRIS, S.A., AHN, S.H., PODOLNY, V., COLLINS, S.R., SCHULDINER, M., CHIN, K., PUNNA, T., THOMPSON, N.J., BOONE, C., EMILI, A., WEISSMAN, J.S., HUGHES, T.R., STRAHL, B.D., GRUNSTEIN, M., GREENBLATT, J.F., BURATOWSKI, S. and KROGAN, N.J. (2005). Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell,* 123:593-605.

KIEFER, J.C. (2007). Epigenetics in Development. *Developmental Dynamics* 236: 1144-1156.

KIM, S.Y., HE, Y., JACOB, Y., NOH, Y.S., MICHAELS, S. and AMASINO, R. (2005). Establishment of the vernalization-responsive, winter-annual habit in *Arabidopsis* requires a putative histone H3 methyl transferase *Plant Cell* 17: 3301–3310.

KONDRASHOV, F.A., ROGOZIN, I.B., WOLF, Y.I. and KOONIN, E.V. (2002). Selection in the evolution of gene duplication. *Genome Biol.* 2: 8.1-8.9.

KRICHEVSKY, A., GUTGARTS, H., KOZLOVSKY, S.V., TZFIRA, T., SUTTON, A., STERNGLANZ, R., MANDEL, G. and CITOVSKY, V. (2007). C2H2 zinc finger-SET histone methyltransferase is a plant-specific chromatin modifier. *Dev Biol* 303:259-269.

KOUZARIDES, T. (2007). Chromatin modifications and their function. *Cell* 128: 693-705.

LI, H., ILIN, S., WANG, W., DUNCAN, E.M., WYSOCKA, J., ALLIS, C.D. and PATEL, D.J. (2006). Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* 442: 91-95.

LI, J., BRADER, G. and PALVA, E.T. (2004). The WRKY70 transcription factor: a node of convergence for jasmonate-mediated and salicylate-mediated signals in plant defense. *Plant Cell* 16: 319-331.

LIU, F., QUESADA, V., CREVILLEN, P., BÃURLE, I., SWIEZEWSKI, S. and DEAN, C. (2007). The Arabidopsis RNA-binding protein FCA requires a lysine-specific demethylase 1 homolog to downregulate FLC. *Mol Cell* 28: 398-407.

LOIDL, P. (2004). A plant dialect of the histone language. *Trends Plant Sci.* 9: 84-90.

MAKAREVICH, G., LEROY, O., AKINCI, U., SCHUBERT, D., CLARENZ, O., GOODRICH, J., GROSSNIKLAUS, U. and KOHLER, C. (2006). Different Polycomb group complexes regulate common target genes in *Arabidopsis*. *EMBO Rep* 7: 947-952.

MARCH-DIAZ, R., GARCIA-DOMINGUEZ, M., FLORENCIO, F.J. and REYES, J.C. (2007). SEF, a new protein required for flowering repression in Arabidopsis, interacts with PIE1 and ARP6. *Plant Physiol* 143: 893-901.

MEIJER, HJG., BERRIER, C.P., IURISCI, C., DIVECHA, N., MUSGRAVE, A. and MUNNIK, T. (2001). Hyperosmotic stress induces rapid synthesis of phosphatidyl-D-inositol 3,5-biphosphatein plant cells. *Biochem J* 360: 491-498.

MORRIS, S.A., RAO, B., GARCIA, B.A., HAKE, S., DIAZ, R.L., SHABANOWITZ, J.,, HUNT, D.F., ALLIS, C.D., LIEb, J.D. and STRAHL, B.D. (2007). Identification of histone H3 lysine 36 acetylation as a highly conserved histone modification. *J Biol Chem* 282: 7632-7640.

NAPOLI, C., LEMIEUX, C. and JORGENSEN, R. (1990). Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. *Plant Cell* 2:279-289.

NG, D.W., WANG, T., CHANDRASEKHARAN, M.B., ARAMAYO, R., KERTBUNDIT, S. and HALL, T.C. (2007). Plant SET domain-containing proteins: structure, function and regulation. *Biochim Biophys Acta* 1769: 316-329.

NONCHEV, S. and TSANEVA, I. (2009). From *Chemical Embryology* to Nucleosome Patterning - an interview with Roumen G. Tsanev. *Int. J. Dev. Biol.* 53: (doi: 10.1387/ijdb.082679sn).

PAPP, B. and MUELLER, J. (2006). Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins. *Genes Dev* 20: 2041-2054.

PIEN, S., FLEURY, D., MYLNE, J., CREVILLEN, P., INZE, D., AVRAMOVA, Z., DEAN, C. and GROSSNIKLAUS, U. (2008). Unraveling trithorax functions in plants: ATX1 dynamically regulates the activation of *FLC* via histone H3 lysine 4 trimethylation. *The Plant Cell* 20, 580-588.

PIEN, S. and GROSSNIKLAUS, U. (2007). Polycomb group amd trithorax group proteins in Arabidopsis. *Biochim Biophys Acta* 1769: 375–382.

PIKAARD, C.S. and PONTES, O. (2007). Heterochromatin: condense or excise. *Nat Cell Biol* 9:19-20.

PIRROTTA, V. (1998). Polycombing the genome: PcG, trxG, and chromatin silencing. *Cell* 93: 333-336.

RAYNAUD, C., SOZZANI, R., GLAB, N., DOMENICHINI, S., PERENNES, C., CELLA, R., KONDOROSI, E. and BERGOUNIOUX, C. (2006). Two cell-cycle regulated SET-domain proteins interact with proliferating cell nuclear antigen (PCNA) in Arabidopsis. *Plant J* 47:395-407.

REA, S., EISENHABER, F., O'CARROLL., STRAHL, B., SUN, Z-W., SCHMIDT, M., OPRAVIL, S., MECHTLER, K., PONTIG, C., ALLIS, D.C. and JENUWEIN, T. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406: 593-599.

REYES, J.C. (2006). Chromatin modifiers that control plant development. *Curr Opin Plant Dev* 9: 21-27.

RUSCHE, L.N., KIRCHMAIER., A.L. and RINE, J. (2003). The establishment, inheritance, and function of silenced chromatin in *Saccharomyces cerevisiae. Annu Rev Biochem* 72: 481-516.

RUTHENBURG, A.J., AKKIS., D.C. and WYSOCKA, J. (2007). Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Cell* 25:15-30.

SAHA, A., WITTMEYER, J. and CAIRNS, B.R. (2006). Chromatin remodelling: the industrial revolution of DNA around histones. *Nat Rev Mol Cell Biol* 7: 437-447.

SALEH, A., AL-ABDALLAT, A., NDAMUKONG, I., ALVAREZ-VENEGAS., R. and AVRAMOVA, Z. (2007). ATX1 and CLF Establish a Bivalent Chromatin Mark at the Silenced *AGAMOUS* Locus in *Arabidopsis. Nucl Acids Res* 35:6290-6296.

SALEH A, ALVAREZ-VENEGAS R. and AVRAMOVA Z. (2008a). Dynamic and stable histone H3 methylation patterns at the Arabidopsis FLC and AP1 loci. *Gene* 423: 43-47

SALEH, A., ALVAREZ-VENEGAS., R., YILMAZ, M., LE, O., HOU, G.,, SADDER, M., AL-ABDALLAT, A., XIA, Y., LU, G., LADUNGA, I. and AVRAMOVA, Z. (2008b). The highly similar *ARABIDOPSIS HOMOLOGS OF TRITHORAX ATX1* and *ATX2* encode divergent biochemical functions. *The Plant Cell* 20: 568-579.

SANTOS-ROSA, H., SCHNEIDER, R., BANNISTER, A., SHERRIFF, J., BERNSTEIN, B.E., EMRE, T., SCHREIBER, S.L., MELLOR, J. and KOUZARIDES, T. (2002). Active genes are tri-methylated at K4 of histone H3. *Nature* 419: 407-411.

SCHOTTA, G., LACHNER, M., SARMA, K., EBERT, A., SENGUPTA, R., REUTER, G., REINbERG, D. and JENUWEIN, T. (2004). A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev* 18:1251-1262.

SCHUBERT, D., CLARENZ, O. and GOODRICH, J. (2005). Epigenetic control of plant development by Polycomb-group proteins. *Curr Opin Plant Biol* 8: 553-561.

SCHUBERT, D., PRIMAVESI, L., BISHOPP, A., RObERTS, G., DOONAN, J., JENUWEIN, T. and GOODRICH, J. (2006). Silencing by plant *Polycomb*-group genes requires dispersed trimethylation of histone H3 at lysine 27. *EMBO J* 25: 4638-4649.

SHI, J. and DAWE, R.K. (2006). Partitioning of the maize epigenome by the number of methyl groups on histone H3 lysines 9 and 27. *Genetics* 173:1571-83.

SPARMANN, A. and VAN, LOHUIZEN, M. (2006). Polycomb silencers control cell fate, development and cancer. *Nat Rev Cancer* 6:846-856.

SPILLANE, C., SCHMID, K.J., LAOUEILLE-DUPRAT, S., PIEN, S., ESCOBAR-RESTREPO, J-M., BAROUX, C., GAGLIARDINI, V., PAGE, D.R., WOLFE, K.H. and GROSSNIKLAUS, U. (2007). Positive Darwinian selection at the imprinted *MEDEA* locus in plants. *Nature* 448: 349-352.

SPRINGER, N.M., NAPOLI, C.A., SELINGER, D.A., PANDEY, R., CONE, K.C., CHANDLER, V.L., KAEPPLER, H.F. and KAEPPLER, S.M. (2003). Comparative analysis of SET domain proteins in maize and Arabidopsis reveals multiple duplications preceding the divergence of monocots and dicots. *Plant Physiol* 132: 907-925.

STASSEN, M.J., BAILEY, D., NELSON, S., CHINWALLA, V. and HARTE, P.J. (1995). The Drosophila trithorax proteins contain a novel variant of the nuclear receptor type DNA binding domain and an ancient conserved motif found in other chromosomal proteins. *Mech Dev*, 52:209-223.

STOKES, T.L., KUNKEL, B.N. and RICHARDS, E.J. (2002). Epigenetic variation in *Arabidopsis* disease resistance. *Genes Dev* 16: 171-182.

STEWARD, M.M., LEE, J-S., O'DONOVAN, A., WYATT, M., BERNSTEIN, B.E. and SHILATIFARD, A. (2007). Molecular regulation of H3H4 trimethylation by ASH2L, a shared subunit of MLL complexes. *Nature Struct. Mol. Biol.*13, 852-854.

SRIDHAR, V.V., KAPOOR, A., ZHANG, K., ZHU, J., ZHOU, T., HASEGAWA, P.M., BRESSAN, R.A. and ZHU, J.K. (2007). Control of DNA methylation and heterochromatic silencing by histone H2B deubiquitination. *Nature* 447: 735-738.

STEVENSON, J.M., PERERA, I.Y., HEILMANN, I., PERSON, S. and BOSS, W. (2000). Inositol signaling and plant growth. *Trends Plant Sci* 5: 252-258.

TURCK, F., ROUDIER, F., FARRONA, S., MARTIN-MAGNIETTE, M.L., GUILLAUME, E., BUISINE, N., GAGNOT, S., MARTIENSSEN, R.A., COUPLAND, G. and COLOT, V. (2007). TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27.*PLoS Genet.* 3(6):e86.

VAN, DIJK, K., MARLEY, K.E., JEONG, R., XU, J., HESSON, J., CERNY, R.L., WATERBORG, J.H. and CERUTTI, H. (2005). Monomethyl histone H3 lysine 4 as an epigenetic mark for silenced euchromatin in Chlamydomonas. *Plant Cell* 17: 2439-2453.

VEERAPPAN, C.S., AVRAMOVA., Z. and MORIYAMA, E.N. (2008). Evolution of SET-domain protein families in the unicellular and multicellular Ascomycota fungi. *BMC Evol Biol* 8: 190-210.

WANG, X. (2004). Lipid signaling. *Curr Opin Plant Biol* 7: 329-336.

WANG, X.,, ZHANG, Y., MA, Q., ZHANG, Z., XUE, Y., BAO, S. and CHONG, K. (2007). SKB1-mediated symmetric dimethylation of histone H4R3 controls flowering time in *Arabidopsis. EMBO J* 26:19341941.

WOOD, C.C., ROBERTSON, M., TANNER, G., PEACOCK, W.J., DENNIS, E.S. and HELLIWELL, C.A. (2006). The Arabidopsis thaliana vernalization response requires a polycomb-like protein complex that also includes VERNALIZATION INSENSITIVE 3. *Proc Natl Acad Sci USA* 103*:* 14631-14636.

WYSOCKA, J., SWIGUT, T., XIAO, H., MILNE, T.A., KWON, S.Y., LANDRY, J., KAUER, M., TACKETT, A.J., CHAIT, B.T., BADENHORST, P., WU, C. and ALLIS, C.D. (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature*, 442: 86-90.

WYSOCKA, J., SWIGUT., T., MILNE., T., DOU., Y., ZHANG., X., BURLINGAME., A., ROEDER., R., BRIVANLOU., A. and ALLIS. and D. (2005). WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* 121: 859-872.

XIA, Z.B., ANDERSON, M., DIAZ, M.O. and ZELEZNIK-LE, N.J. (2003). MLL repression domain interacts with histone deacetylases, the polycomb group proteins HPC2 and BMI-1, and the corepressor C-terminal-binding protein. *Proc Natl Acad Sci USA* 100: 8342-8347.

ZHANG, X., CLARENZ, O., COKUS, S., BERNATAVICHUTE, Y.V., PELLEGRINI, M., GOODRICH, J. and JACOBSEN, S.E. (2007a). Whole-genome analysis of histone H3 lysine 27 trimethylation in *Arabidopsis. PLoS Biol* 5(5):e129.

ZHANG, K., SRIDHAR, V.V., ZHU, J., KAPOOR, A. and ZHU, J-K. (2007b). Distinctive core histone post-translational modification patterns in Arabidopsis thaliana. *PloS ONE* 2(11): e1210.doi:10.1371.

ZHANG, X., YANG, Z., KHAN, S.I., HORTON, J.R., TAMARU, H., SELKER, E.U. and CHENG, X. (2003).Structural basis for the product specificity of histone lysine methyltransferases *Mol Cell* 12: 177–185.

ZHAO, Z., YU, Y., MEYER, D., WU, C. and SHEN, W.H. (2005). Prevention of early flowering by expression of *FLOWERING LOCUS C* requires methylation of histone H3 K36 *Nature Cell Biol* 7:1156–1160.